Methods and means for modulating cellulose biosynthesis in fiber producing plants.

[0001] Field of the invention.

[0002] The invention relates to the field of agricultural biotechnology. More specifically, the invention provides novel genes involved in cellulose biosynthesis and methods using such genes to modulate cellulose biosynthesis in fiber-producing plants such as cotton. The invention also provides methods for identifying and isolating alleles of these genes in a population of fiber producing plants that correlate with the quality of the produced fibers.

[0003] Background

[0004] Cellulose is the major structural polysaccharide of higher plant cell walls. Chains of β-1,4-linked glucosyl residues assemble soon after synthesis to form rigid, chemically resistant microfibrils. Their mechanical properties together with their orientation in the wall influence the relative expansion of cells in different directions and determine many of the final mechanical properties of mature cells and organs. These mechanical properties are of great importance for wood, paper, textile and chemical industries.

[0005] Much of the high quality fiber for the textile industry is provided for by cotton. About 90% of cotton grown worldwide is *Gossypium hirsutum* L., whereas *Gossypium barbadense* accounts for about 8%.

[0006] Several genes involved in cellulose biosynthesis have already been identified by mutational analysis in a number of plants. Mutants of *Arabidopsis thaliana* show that *in vivo* cellulose synthesis requires the activity of members of the At*Ces*A gene family encoding glycosyltransferases (Arioli et al., 1998; Taylor et al., 1999; Fagard et al., 2000;Taylor et al., 2000; Scheible et al., 2001; Burn et al., 2002a; Desprez et al., 2002), of the AtKOR1 gene (At5g49720) encoding a membrane-associated endo-1,4-ß-D-glucanase (Nicol et al., 1998; Zuo et al., 2000; Lane et al., 2001; Sato et al., 2001), of KOBITO1 encoding a plasma membrane protein of unknown function (Pagant et al., 2002) and of genes encoding enzymes in the N-glycosylation/quality control pathway in the ER (Lukowitz et al., 2001; Burn et al., 2002b; Gillmor et al., 2002).

[0007] The function of an endo-1,4-β-D-glucanase in cellulose synthesis remains to be determined but the lack of activity against crystalline cellulose of BnCel16, a related *Brassica*

napus enzyme (Mølhøj et al., 2001), suggests that the enzyme probably cleaves a non-crystalline glucan chain such as a lipid-linked primer or glucan donor (Williamson et al., 2001; Peng et al., 2002). Tomato Cel3 (LeCel3) was the first such membrane-associated endo-1,4-β-D-glucanase identified (Brummell et al., 1997) and antibodies to LeCel3 detected a cotton fiber protein upregulated during herbicide inhibition of cellulose synthesis (Peng et al., 2001). A cotton fiber membrane fraction required Ca^{2+} for in vitro cellulose synthesis activity and, because an exogenous, Ca^{2+}-independent endo-1,4-β-D-glucanase restored cellulose synthesis activity, a cotton orthologue of KOR (GhKOR) was proposed as the endogenous Ca^{2+}-dependent factor (Peng et al., 2002). A truncated form of BnCel16 showed Ca^{2+}-dependence in vitro (Mølhøj et al., 2001).

[0008] Further genetic data point to cellulose synthesis responding to defects in enzymes on the N-glycosylation/quality control pathway. These steps occur in the ER rather than at the plasma membrane and so probably act only indirectly on synthesis through the supply of key glycoproteins to the plasma membrane. N-glycosylation begins when the mannose-rich oligosaccharide Glc3Man9GlcNac2 is assembled on dolichol in the ER membrane and transferred to the Asn residue of a newly synthesized protein containing an Asn-X-Ser or Asn-X-Thr motif (where X is any amino acid except Pro).

[0009] With further processing of the glycoprotein by glucosidases I and II, N-glycosylation intersects with the quality control pathway responsible for ensuring proper folding of newly synthesized proteins (Helenius and Aebi, 2001; Vitale, 2001). Glucosidase I removes the terminal α-l,2-linked glucosyl residue to generate Glc2Man9GlcNac2 and glucosidase II removes the next α-l,3-glucosyl residue. Polypeptides carrying the resultant GlcMan9GlcNac2 specifically bind chaperones (calnexin and calreticulin) and probably other proteins that promote proper folding of newly synthesized proteins. The glycoprotein releases the chaperones when glucosidase II trims of the final Glc residue which is required for chaperone binding. Glycoprotein glucosyltransferase then reattaches one Glc residue to the Man9GlcNAc2 of improperly folded glycoproteins so that they again bind chaperones and have a further opportunity to fold properly. Properly folded proteins, however, cannot be reglucosylated by that enzyme and progress though the secretory pathway for further processing and delivery.

[0010] Defects at several points in this pathway affect cellulose synthesis. Sequence analysis suggests that the potato MALl gene encodes a glucosidase II and antisense suppression reduces

glucosidase II activity (Taylor et al, 2000a).M4LJ antisense plants accumulate less cellulose than controls when grown under field conditions although there is no visible phenotype in glasshouse conditions. The embryo lethal knopf mutant is deficient in glucosidase I and severely deficient in cellulose (Gillmor et al., 2002). Finally the embryo lethal cyt1 mutant is cellulose-deficient from a defect in mannose- 1-phosphate guanylyltransferase, the enzyme generating the UDP-Man required to (amongst other things) assemble the high mannose oligosaccharide that is transferred from dolichol to the nascent protein (Lukowitz et al, 2001). The mutations that affect cellulose synthesis concentrate towards those early steps where the N-glycosylation pathway intersects with the quality control pathway. Quality control, rather than production of mature glycans on critical proteins, seems particularly important since there is no detectable phenotype from a defect in N-acetyl glucosaminyl transferase I that blocks the steps in the Golgi that build mature, N-linked glycans (von Schaewen et al, 1993).

[0011] Baskin et al. 1992 described *Arabidopsis* mutants which show root radial swelling, named rsw1, rsw2 and rsw3. These mutant lines where shown to exhibit a selective reduction in cellulose production (Peng et al. 2000).

[0012] WO98/00549 relates generally to isolated genes which encode polypeptides involved in cellulose biosynthesis in plants and transgenic plants expressing same in sense or antisense orientation, or as ribozymes, co-suppression or gene-targeting molecules. More particularly, this disclosure is directed to a nucleic acid molecule isolated from *Arabidopsis thaliana*, *Oryza sativa*, wheat, barley, maize, *Brassica spp. Gossypium hirsutum* and *Eucalyptus spp*, which encode an enzyme which is important in cellulose biosynthesis, in particular the cellulose synthase enzyme and homologues, analogues and derivatives thereof and uses of same in the production of transgenic plants expressing altered cellulose biosynthetic properties.

[0013] WO 98/50568 discloses the use of a nucleotide sequence coding for an endo-1,4-β-glucanase to inhibit cell growth in a plant. The nucleotide sequence corresponds wholly or partially to the *Arabidopsis* KOR protein sequence, or to a protein sequence the N-terminal end of which has at least 40% identity with the first 107 amino acids of said KOR, or at least 70% identity with the first 107 amino acids of said KOR.

[0014] WO 97/24448 describes recombinant and isolated nucleic acids encoding a plant α-glucosidase enzyme. An antisense nucleotide was also provided as well as the use of both the

isolated or recombinant sequences and the antisense sequences. Uses of the invention include enhancing and reducing expression of alpha-glucosidases and the provision of novel starches.

[0015] WO 00/08175 relates to nucleic acid molecules coding for a protein with the activity of an alpha-glucosidase from a potato. The invention also relates to methods for the production of transgenic plant cells and plants synthesizing modified starch. The invention further relates to vectors and host cells containing the nucleic acid molecules, plant cells and plants obtained according to the methods, starch synthesized by the described plant cells and methods for the production of such starch.

[0016] WO 98/39455 discloses a gene and enzyme participating in the synthesis of cellulose by microorganisms. A specific gene encoding a cellulase, cellulose synthase complex and alpha-glucosidase are described.

[0017] WO9818949 and US6271443 provide two plant cDNA clones that are homologs of the bacterial CelA genes that encode the catalytic subunit of cellulose synthase, derived from cotton (*Gossypium hirsutum*). Also provided are genomic promoter regions to these encoding regions to cellulose synthase. Methods for using cellulose synthase in cotton fiber and wood quality modification are also provided.

[0018] The prior art remains however deficient in providing alternatives to the known genes involved in cellulose biosynthesis and does not disclose the nucleotide sequence of the wild type gene involved in cellulose biosynthesis and mutated in the *rsw3* mutant Arabidopsis line. Also, the prior art does not disclose the cotton homologues genes of *RSW2* or *RSW3* involved in cellulose biosynthesis from cotton.

[0019] These and other problems have been solved as set forth hereinafter in the different embodiments and claims of the invention.

[0020] Summary of the invention

[0021] It is one object of the invention to provide a method for increasing cellulose biosynthesis e.g. in lint fiber, in fiber-producing plants, such as cotton plants, comprising the steps of

(a) providing cells of said fiber-producing plant with a chimeric gene comprising the following operably linked DNA fragments

i) a promoter expressible in said cell of said plant, such as a constitutive promoter, a fiber specific promoter or an expansin promoter;

ii) a DNA region coding for the protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8 (or a variant of that protein having the same enzymatic activity), such as the nucleotide sequence of SEQ ID No 1 from the nucleotide at position to the nucleotide at position 1986 or SEQ ID No. 2 from the nucleotide position 47 to the nucleotide at position 1906 or SEQ ID No 3 or SEQ ID No 4 from the nucleotide position 2 to the nucleotide at position 1576 or SEQ ID No. 9;

iii) a 3' region involved in transcription termination and polyadenylation.

[0022] It is another object of the invention to provide a method for decreasing cellulose biosynthesis in fiber-producing plants, for example in cotton plants, e.g. in fuzz fiber, comprising the step of providing cells of said fiber-producing plant with a chimeric gene capable of reducing the expression of a gene endogenous to said fiber-producing plant, wherein said endogenous gene codes for a protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8 or a variant thereof, said variant having the same enzymatic activity. The introduced chimeric gene may comprise a nucleotide sequence of 21 contiguous nucleotides selected from a nucleotide sequence which codes for a protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8, such as the nucleotide sequence of SEQ ID No 1 or SEQ ID No. 2 or SEQ ID No 3 or SEQ ID No 4 or SEQ ID No. 9, or the complement thereof, operably linked to a plant expressible promoter, such as a constitutive promoter or a fuzz fiber specific promoter and a 3' region involved in transcription termination and polyadenylation. The chimeric gene may also comprise a first nucleotide sequence of 21 contiguous nucleotides selected from a nucleotide sequence which codes for a protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8, such as the nucleotide sequence of SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3 and SEQ ID No. 4 or SEQ ID No. 9, and a second nucleotide sequence complementary to the first nucleotide sequence, operably linked to a plant expressible promoter and a 3' region involved in transcription termination and polyadenylation such that upon transcription of said chimeric gene, a RNA is formed which can form a double stranded RNA region between said first and said second nucleotide sequence.

[0023] The invention further relates to a chimeric gene for increasing cellulose biosynthesis in fiber-producing plants, e.g. in cotton plants, comprising the following operably linked DNA fragments: a promoter expressible in said cell of said plant such as a constitutive promoter, a (lint)-fiber specific promoter or an expansin promoter; a DNA region coding for the protein comprising the amino acid sequence of SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8 or a variant thereof, said variant having the same enzymatic activity, such as the nucleotide sequence of SEQ ID No. 1 from the nucleotide at position 121 to the nucleotide at position 1986 or SEQ ID No 2 from the nucleotide at position 47 to the nucleotide at position 1906 or SEQ ID No 3 or SEQ ID No 4 from the nucleotide at position 2 to the nucleotide at position 1576 or SEQ ID No. 9; and a 3'end region involved in transcription termination and polyadenylation.

[0024] The invention also relates to a chimeric gene for decreasing cellulose biosynthesis in fiber-producing plants, e.g. in cotton plants, comprising a nucleotide sequence of 21 contiguous nucleotides selected from a nucleotide sequence which codes for a protein comprising the amino acid sequence of SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8, such as the nucleotide sequence of SEQ ID No. 2, SEQ ID No. 3 or SEQ ID No. 4 or SEQ ID No. 9, or the complement thereof, operably linked to a plant expressible promoter and a 3' region involved in transcription termination and polyadenylation.

[0025] The invention further relates to a chimeric gene for decreasing cellulose biosynthesis in fiber-producing plants, e.g. in cotton plants, comprising a first nucleotide sequence of 21 contiguous nucleotides selected from a nucleotide sequence which codes for a protein comprising the amino acid sequence of SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8, and a second nucleotide sequence complementary to said first nucleotide sequence, operably linked to a plant expressible promoter and a 3' region involved in transcription termination and polyadenylation such that upon transcription of said chimeric gene, a RNA is formed which can form a double stranded RNA region between said first and said second nucleotide sequence.

[0026] It is yet another object of the invention to provide plant cells and plants comprising the chimeric genes of the invention as well as seeds of such plants comprising the chimeric genes of the invention.

[0027] The invention thus relates to the use of a chimeric gene according to the invention to modulate cellulose biosynthesis and fiber quality in a fiber producing plant, such as cotton.

[0028] It is also an object of the invention to provide a method for identifying allelic variations of the genes encoding proteins involved in cellulose biosynthesis in a population of different genotypes or varieties of a particular plant species, for example a fiber-producing plant species, which are correlated either alone or in combination with the quantity and/or quality of cellulose production, and fiber production comprising the steps of:

a) providing a population of different varieties or genotypes of a particular plant species or interbreeding plant species comprising different allelic forms of the nucleotide sequences encoding proteins comprising the amino acid sequences of SEQ ID No 5, 6, 7 or 8;
b) determining parameters related to fiber production and/or cellulose biosynthesis for each individual of the population;
c) determining the presence (or absence) of a particular allelic form of the nucleotide sequences encoding proteins comprising the amino acid sequences of SEQ ID No 5, 6, 7 or 8 for each individual of the population; and
d) correlating the occurrence of particular fiber or cellulose parameters with the presence of a particular allelic form of the mentioned nucleotide sequence or a particular combination of such allelic forms.

[0029] Brief Description of the figures

[0030] Figure 1. ClustalW alignment of proteins GhKOR (SEQ ID No 6), LeCel3 (Accession number T07612) and AtKOR1 (Accession number At5g49720; SEQ ID No 5) and BnCel16 (Accession number CAB51903). Features highlighted are: polarized targeting motifs implicated in targeting to the cell plate (Zuo et al., 2000); a putative transmembrane region near the N-terminus (transmembrane); four of the conserved residues potentially involved in catalysis (Asp-198, Asp-201, His-516 and E-555; labeled o) and representing part of the strong similarity to family 9 glycoside hydrolases; a C-terminal region rich in Pro and characteristic of membrane-bound members of the endo-1,4-β-glucanase family; 8 putative N-glycosylation sites (Asn-X-Ser/Thr; labeled G1 to G8).

[0031] Figure 2. Complementation of rsw2-1 by transformation with GhKOR1 cDNA (SEQ ID No 2), operably linked to the CaMV35S promoter. (A) Roots of rsw2-1 swell after exposure to 29°C for 2 d but wild type (Co) and complemented plants containing either AtKOR1 or GhKOR do not. (B) Mature stems of two plants each of rsw2-1 (left), wild type and rsw2-1 expressing

GhKOR. Photograph of plants grown in pots at 21°C, until bolting began, at which time bolts were cut off and plants transferred to 29°C for bolts to regrow.

[0032] Figure 3. Mutations in the gene encoding glucosidase II cause radial swelling. (a) Complementation of root radial swelling in *rsw3* transformed with the 5.8 kB fragment amplified from the wild-type genome. Columbia wild type (left), rsw3 (center) and a kanamycin-resistant T1 seedling of *rsw3* transformed with a genomic copy of the glucosidase II gene (right). The wild type gene suppresses radial swelling. All plants were transferred to 30°C for 2 d prior to photographing. (b) The rsw3 mutation is allelic to the insertional mutant 5GT5691 which contains a Ds element in the first exon of the glucosidase II gene. Columbia wild type (left), *rsw3* (center) and a heterozygous F1 plant from crossing 5GT5691 with *rsw3*. The F1 heterozygote and the *rsw3* homozygote show temperature-induced radial swelling. All plants were transferred to 30°C for 2 d prior to photographing.

[0033] Figure 4. Alignment of the Aglu-3/RSW3 sequence (Genbank NP_201189) with the sequences of ER-resident glucosidase II enzymes from potato (Accession number T07391), mouse (NP_032086) and fission yeast (CAB65603). The clade 2 of Monroe *et al.* (1999) are shown to demonstrate the high conservation. They include several residues implicated in catalysis (Asp 512 and Asp 617; *). The site of the *rsw3*-1 mutation (Ser599•) is close to these consensus sequences and is conserved in these and other glucosidase II sequences. Predicted N-terminal signal sequences are boxed. No HDEL ER-retention sequences occur at the C-terminus.

[0034] Figure 5. Alignments of the proposed β-subunits of *Arabidopsis* (At5g56360) and rice (our amendment of BAA88186) with the β-subunits of glucosidase II from mouse (AAC53183) and fission yeast (BAA13906). Note the predicted N-terminal signal sequences (boxed), C-terminal H/VDEL ER-retention signals and the mannose-receptor homology region (MHR) near the N-terminus. The 6 cysteines within the MHR (four only in yeast) are numbered, and the R and Y residues implicated in substrate-binding (•) and the substrate recognition loop between cysteines 5 and 6 are marked. Elsewhere in the sequence, note the relatively high level of similarity in the N- and C-terminal domains and the much lower similarity and plant-specific inserts in the central region.

[0035] Figure 6. mRNA for both the α-subunit (a) and the β-subunit (b) occurs in all *Arabidopsis* tissues tested. RT-PCR using mRNA from root (lane 1), whole rosette leaves (2), leaf blades (3), mature stem tissue (4), cauline leaves (5), flower buds (6), flowers (7), siliques

(8), dark grown hypocotyls (9). (The presence of the β-subunit in dark grown hypocotyls was demonstrated in another experiment).

[0036] Figure 7. Morphology of *rsw3*.

(a) Root system of a seedling showing that lateral roots extend some distance before swelling and stopping elongation. Plants grown 5d at 21°C and 6 d at 30°C. Scale bar = 2 mm.

(b) Continued root growth gives a dense, highly branched root system and a dense mass of very small leaves on a plant grown for 21 d at 30°C. Scale bar = 5mm.

(c) Hypocotyls grown in the dark for 3 d at 21°C and 3 d at 30°C. From the left: wildtype, *rsw1-1, rsw2-1, rsw3, rsw1-1rsw2-1, rsw1-1rsw3*. The *rsw3* effect on the hypocotyl is weak compared to that of the other single mutants and *rsw1-1rsw3* is weaker than *rsw1-1rsw2-1*. Scale bar = 5 mm.

(d) Light micrograph of *rsw3* grown on agar for 35 d at 30°C. Tiny inflorescences with flower buds of near normal size (top right and bottom left) emerge from several of the rosettes. Scale bar = 5 mm.

(e) Scanning electron micrograph of *rsw3* plant grown for 21 d at 30°C and showing the presence of multiple rosettes. Scale bar = 1 mm.

(f) Detail of the ringed area in (e) showing the very complex arrangement of the minute leaves, many of which carry trichomes of approximately normal size and morphology. Scale bar = 200 μm.

(g) Scanning electron micrograph of the surface of a wild type leaf on a plant grown for 10 d at 30°C. Note the clearly defined cell boundaries, stomata and trichomes.

(h) The surface of an *rsw3* leaf showing much less clear outlines to the pavement cells, an apparently collapsed trichome (CT) on top of its ring of subsidiary cells and many stomata with their guard cells protruding above the leaf surface. Scale bar for (g) and (h) = 100 μm.

[0037] Figure 8. Growth of the stem and reproductive development in *rsw3*.

(a and b) Kinetics of secondary stem elongation in Columbia wild type, *rsw3*, *rsw1* and the *rsw1rsw3* double mutant at 21°C (a) and 30°C (b). All plants were grown at 21°C until stems began to emerge. These were cut off and re-growth of secondary bolts followed at the indicated temperature. Single mutants show very little difference from wild type at 21°C although the double mutant elongates more slowly and reaches a significantly shorter final height. The final heights reached at 30°C differ widely as do the trajectories by which they are reached. *rsw1*

elongates more slowly but elongation continues for at least as long as it does in wild type. *rsw3* elongates almost as rapidly as wild type for 4 d but then ceases elongation by about day 6. The *rsw1rsw3* double mutant elongates less rapidly and ceases elongation at about day 5.
(c and d). Light micrographs showing well spaced flowers in wild type (c) and the clustered flowers on *rsw3* (d) with its early cessation of elongation.
(e and f) Cryoscanning electron micrographs showing flower buds of wild type (e) and *rsw3* (f) that are of similar sizes but open prematurely in *rsw3*. Note the immature state of the stigma (St) and the irregular shapes of the cells on the sepals (Se) in *rsw3*. Bar for (e) and (f) = 200 μm.
(g and h) Cryo-scanning electron micrographs showing imbibed seed of *rsw3* that developed on plants held at 21°C (g) and 3O°C (h). The 30°C seed is shrunken and lacks the clear cellular pattern of the 21°C seed.
(i-n) Light micrographs of imbibed seed stained with ruthenium red to show a surface coat of mucilage. Wild type (i,j), *rsw1* (k,l), *rsw3* (m,n). Seed in i, k, m developed on plants at 21°C, seed in j, l, n developed on plants at 30°C. Mucilage is secreted normally by *rsw1* (l) and wild type (j) at 30°C but not by *rsw3*(n).

[0038] Detailed description

[0039] The invention is based on the identification of the wild type gene which has been mutated in *Arabidopsis* mutant *rsw3*, and elucidation of its function. The inventors have also identified the cotton genes corresponding to the genes mutated in *rsw2* and *rsw3 Arabidopsis* mutants. These cotton genes are implicated in cellulose production.

[0040] In one embodiment the invention thus relates to a method for increasing the production of cellulose in a plant comprising the steps of providing cells of the plant with a chimeric gene comprising a plant-expressible promoter operably linked to a DNA region coding for a protein comprising the amino acid sequence of SEQ ID No 5, SEQ ID No. 6, SEQ ID No 7 or SEQ ID No 8 or a variant thereof having similar activity as the mentioned proteins, and a 3' region involved in transcription termination and polyadenylation. The plants may be fiber-producing plants such as cotton, and the increased cellulose production may result in a larger production of cotton fibers, e.g. cotton lint fibers, or in cotton fibers with altered or increased length, or altered quality such as improved tensile strength.

[0041] As used herein, "chimeric gene" or "chimeric nucleic acid" refers to any gene or any nucleic acid, which is not normally found in a particular eukaryotic species or, alternatively, any gene in which the promoter is not associated in nature with part or all of the transcribed DNA region or with at least one other regulatory region of the gene.

[0042] As used herein, the term "promoter" denotes any DNA which is recognized and bound (directly or indirectly) by a DNA-dependent RNA-polymerase during initiation of transcription. A promoter includes the transcription initiation site, and binding sites for transcription initiation factors and RNA polymerase, and can comprise various other sites (e.g., enhancers), at which gene expression regulatory proteins may bind. The term "regulatory region", as used herein, means any DNA, that is involved in driving transcription and controlling (i.e., regulating) the timing and level of transcription of a given DNA sequence, such as a DNA coding for a protein or polypeptide. For example, a 5' regulatory region (or "promoter region") is a DNA sequence located upstream (i.e., 5') of a coding sequence and which comprises the promoter and the 5'-untranslated leader sequence. A 3' regulatory region is a DNA sequence located downstream (i.e., 3') of the coding sequence and which comprises suitable transcription termination (and/or regulation) signals, including one or more polyadenylation signals.

[0043] In one embodiment of the invention the promoter is a constitutive promoter. In another embodiment of the invention, the promoter activity is enhanced by external or internal stimuli (inducible promoter), such as but not limited to hormones, chemical compounds, mechanical impulses, abiotic or biotic stress conditions. The activity of the promoter may also be regulated in a temporal or spatial manner (tissue-specific promoters; developmentally regulated promoters).

[0044] In a particular embodiment of the invention, the promoter is a plant-expressible promoter. As used herein, the term "plant-expressible promoter" means a DNA sequence which is capable of controlling (initiating) transcription in a plant cell. This includes any promoter of plant origin, but also any promoter of non-plant origin which is capable of directing transcription in a plant cell, i.e., certain promoters of viral or bacterial origin such as the CaMV35S (Hapster et al., 1988), the subterranean clover virus promoter No 4 or No 7 (WO9606932), or T-DNA gene promoters but also tissue-specific or organ-specific promoters including but not limited to seed-specific promoters (e.g., WO89/03887), organ-primordia specific promoters (An et al., 1996), stem-specific promoters (Keller et al., 1988), leaf specific promoters (Hudspeth et al.,

1989), mesophyl-specific promoters (such as the light-inducible Rubisco promoters), root-specific promoters (Keller et al.,1989), tuber-specific promoters (Keil et al., 1989), vascular tissue specific promoters (Peleman et al., 1989), stamen-selective promoters (WO 89/10396, WO 92/13956), and the like.

[0045] Suitable plant-expressible promoters include the fiber specific and/or secondary cell wall specific promoters which can be isolated according to the teaching of WO 98/18949, WO98/00549 or US5932713. Also suitable are the promoters disclosed in WO98/18949 or US 6,271,443. Cotton lint-fiber specific promoters are also suitable.

[0046] In one embodiment of the above mentioned methods, the DNA region coding for a protein comprising the amino acid sequence of SEQ ID No 5, SEQ ID No 6, SEQ ID No 7 or SEQ ID No 8 comprises the nucleotide sequence of SEQ ID No 1 from nucleotide 121 to nucleotide 1986, SEQ ID No 2 from nucleotide 47 to nucleotide 1906, SEQ ID No. 3 or SEQ ID No. 4 from nucleotide 2 to nucleotide 1576 or SEQ ID No. 9.

[0047] In another embodiment of the above mentioned methods, the DNA region codes for a variant of the proteins comprising the amino acid sequence of SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7 or SEQ ID No. 8. As used herein, "variant" proteins refer to proteins wherein one or more amino acids are different from the corresponding position in the proteins having the amino acid sequence of SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7 or SEQ ID No. 8, by substitution, deletion, insertion; and which have at least one of the functions of the proteins encoded by SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7 or SEQ ID No. 8 such as e.g. the same enzymatic or catalytic activity. Methods to derive variants such as site-specific mutagenesis methods are well known in the art, as well as assays to identify the enzymatic activity encoded by the variant sequences. Suitable substitutions include, but are not limited to, so-called conservative substitutions in which one amino acid residue in a polypeptide is replaced with another naturally occurring amino acid of similar chemical character, for example Gly⇔Ala, Val⇔Ile⇔Leu, Asp⇔Glu, Lys⇔Arg, Asn⇔Gln or Phe⇔Trp⇔Tyr.

[0048] Allelic forms of the nucleotide sequences which may encode variant proteins, according to the specification may be identified by hybridization of libraries, under stringent conditions, such as cDNA or genomic libraries of a different varieties or plant lines, e.g. cotton varieties and plant lines. Nucleotide sequences which hybridize under stringent conditions to nucleotide sequences encoding the amino acid sequence of SEQ ID 5, 6, 7 or 8 or to the nucleotide

sequence of SEQ ID 1, 2, 3, 4 or 9, or a sufficiently large part thereof (e.g., at least about 25 contiguous nucleotides, at least about 50 contiguous nucleotides, or at least about 100 contiguous nucleotides) and which encode a functional protein that can complement at least one function, and may complement all of the affected functions, in the *rsw2* or *rsw3* mutant line in *Arabidopsis* are functional equivalents of the above mentioned coding regions. Such nucleotides may also be identified and isolated using e.g. polymerase chain reaction amplification using an appropriate pair of oligonucleotides having at least about 25 contiguous nucleotides, at least about 50 contiguous nucleotides, or at least about 100 contiguous nucleotides of the nucleotide of SEQ ID No 1, SEQ ID No 2, SEQ ID No. 3 , SEQ ID No 4 or SEQ ID No. 9.

[0049] "Stringent hybridization conditions" as used herein mean that hybridization will generally occur if there is at least 95%, or at least 97%, sequence identity between the probe and the target sequence. Examples of stringent hybridization conditions are overnight incubation in a solution comprising 50% formamide, 5× SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5× Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared carrier DNA such as salmon sperm DNA, followed by washing the hybridization support in 0.1× SSC at approximately 65°C. Other hybridization and wash conditions are well known and are exemplified in Sambrook et al, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, NY (1989), particularly chapter 11.

[0050] As another aspect of the invention, the identified genes may be used to decrease cellulose biosynthesis in plants such as fiber-producing plants, e.g. cotton. Thus, in another embodiment of the invention, a method is provided to decrease cellulose biosynthesis in plants such as fiber-producing plants, e.g. in cotton plants, comprising the step of providing cells of said fiber-producing plant with a chimeric gene capable of reducing the expression of a gene endogenous to said fiber-producing plant, wherein said endogenous gene codes for a protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8 or a variant thereof, said variant having the same functional or enzymatic activity.

[0051] In one embodiment of this method of the invention, a chimeric gene is provided to cells of the plant, wherein the chimeric gene comprises a nucleotide sequence of 21 contiguous nucleotides selected from a nucleotide sequence which codes for a protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8, such as a nucleotide sequence of 21 contiguous nucleotides selected from the nucleotide sequences of SEQ

ID No. 1 or SEQ ID No 2 or SEQ ID No 3 or SEQ ID No 4 or SEQ ID No. 9 operably linked to a plant expressible promoter and a 3' region involved in transcription termination and polyadenylation (so-called "sense" RNA mediated gene silencing). In another embodiment of this method of the invention, a chimeric gene is provided to cells of the plant, wherein the chimeric gene comprises a nucleotide sequence of 21 contiguous nucleotides selected from the complement of a nucleotide sequence which codes for a protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8, such as a nucleotide sequence of 21 contiguous nucleotides selected from the complement of the nucleotide sequences of SEQ ID No. 1 or SEQ ID No 2 or SEQ ID No 3 or SEQ ID No 4 or SEQ ID No. 9 operably linked to a plant expressible promoter and a 3' region involved in transcription termination and polyadenylation (so-called "antisense" RNA mediated gene silencing).

[0052] The length of the antisense or sense nucleotide sequence may vary from about 21 nucleotides (nt), up to a length equaling the length (in nucleotides) of the target nucleic acid. The total length of the antisense or sense nucleotide sequence may be at least about 50 nt, 100 nt, 150 nt, 200 nt, or 500 nt long. It is expected that there is no upper limit to the total length of the antisense nucleotide or sense nucleotide sequence, other than the total length of the target nucleic acid. However for practical reason (such as, e.g., stability of the chimeric genes) the length of the antisense or sense nucleotide sequence may be limited to 5000 nt, to 2500 nt, or even to about 1000 nt.

[0053] It will be appreciated that the longer the total length of the antisense or sense nucleotide sequence is, the less stringent the requirements for sequence identity between the total antisense or sense nucleotide sequence and the corresponding sequence in the target gene or the complement thereof become. In one embodiment, the total antisense nucleotide sequence will have a sequence identity of at least about 75% with the complement corresponding target sequence; alternatively, at least about 80 %, at least about 85%, about 90%, about 95%, about 100%, or is identical to complement of the corresponding part of the target nucleic acid. In one embodiment, the antisense or sense nucleotide sequence will include a sequence of about 20-21 nt with 100% sequence identity to the corresponding part of the target nucleic acid or the complement thereof. For calculating the sequence identity and designing the corresponding

antisense or sense sequence, the number of gaps may be minimized, particularly for the shorter antisense or sense sequences.

[0054] For the purpose of this invention, the "sequence identity" of two related nucleotide or amino acid sequences, expressed as a percentage, refers to the number of positions in the two optimally aligned sequences which have identical residues (×100) divided by the number of positions compared. A gap, i.e., a position in an alignment where a residue is present in one sequence but not in the other, is regarded as a position with non-identical residues. The alignment of the two sequences may be performed by the Needleman and Wunsch algorithm (Needleman and Wunsch, 1970) Computer-assisted sequence alignment, can be conveniently performed using standard software program such as GAP which is part of the Wisconsin Package Version 10.1 (Genetics Computer Group, Madison, Wisconsin, USA) using the default scoring matrix with a gap creation penalty of 50 and a gap extension penalty of 3.

[0055] Another embodiment of the invention, relates to a method for reducing the expression of endogenous genes of said fiber-producing plant, wherein said endogenous gene codes for a protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8 or a variant thereof using DNA regions, under the control of a plant-expressible promoter, which when transcribed result in so-called double stranded RNA molecules, comprising both sense and antisense sequences which are capable of forming a double stranded RNA molecule as described in WO 99/53050 (herein entirely incorporated by reference).

[0056] Thus, in one embodiment of the invention, a chimeric gene may be provided to a plant cell comprising a plant expressible promoter operably linked to a DNA region, whereby that DNA region comprises a part of coding region comprising at least 20 or 21consecutive nucleotides from the coding region of a nucleic acid encoding a protein with the amino acid sequence of SEQ ID Nos 5, 6, 7 or 8 (the so-called sense part) as well as a DNA sequence that comprises at least the complementary DNA sequence of at least 20 or 21 nucleotides of the sense part, but which may be completely complementary to the sense part (the so-called antisense part). The chimeric gene may comprise additional regions, such as a transcription termination and polyadenylation region functional in plants. When transcribed an RNA can be produced which may form a double stranded RNA stem between the complementary parts of the sense and antisense region. A spacer region may be present between the sense and antisense nucleotide

sequence. The chimeric gene may further comprise an intron sequence, which may be located in the spacer region.

[0057] In yet another embodiment of the invention, the chimeric gene used to reduce the expression of a gene endogenous to said fiber-producing plant, wherein said endogenous gene codes for a protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8 or a variant thereof, said variant having the same functional or enzymatic activity, encodes a ribozyme which recognizes and cleaves RNA having the nucleotide sequence of an RNA coding for a protein comprising the amino acid sequence of SEQ ID No. 5 or SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8 or a variant thereof. In another embodiment, the ribozyme recognizes and cleaves RNA having the nucleotide sequence of an RNA comprising the nucleotide sequence of SEQ ID 1, 2, 3 or 4. Methods for designing and using ribozymes have been described by Haseloff and Gerlach (1988) and are contained *i.a.* in WO 89/05852.

[0058] It will be clear that whenever nucleotide sequences of RNA molecules are defined by reference to nucleotide sequence of corresponding DNA molecules, the thymine (T) in the nucleotide sequence should be replaced by uracil (U). Whether reference is made to RNA or DNA molecules will be clear from the context of the application. In yet another embodiment of the invention, nucleic acids (either DNA or RNA molecules) are provided which can be used to alter cellulose biosynthesis in plants. Thus the invention provides chimeric genes (DNA molecule) which comprise the following operably linked DNA fragments

i) a promoter expressible in said cell of said plant;

ii) a DNA region comprising a nucleotide sequence of at least 21 nucleotides selected from a nucleotide sequence coding for the protein comprising the amino acid sequence of SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8 (or a variant of that protein having the same enzymatic activity), such as the nucleotide sequence of SEQ ID Nos 1, 2, 3 , 4 or 9 ; and/or

iii) a DNA region and comprising a nucleotide sequence of at least 21 nucleotides selected from the complement of a nucleotide sequence coding for the protein comprising the amino acid sequence of SEQ ID No 6 or SEQ ID No 7 or SEQ ID No 8 or a variant thereof, said variant having the same enzymatic activity, such as the nucleotide sequence of SEQ ID Nos 1, 2, 3, 4 or 9; and

iv) a 3'end region involved in transcription termination and polyadenylation.

[0059] Also provided are RNA molecules that can be obtained from the chimeric genes according to the invention. Such RNA molecules can be produced by *in vivo* or *in vitro* transcription of the chimeric genes. They can also be obtained through *in vitro* transcription of chimeric genes, wherein the transcribed region is under control of a promoter recognized by single subunit RNA polymerases from bacteriophages such as SP6, T3 or T7. Alternatively, the RNA molecules may be synthesized *in vitro* using procedures well known in the art. Also chemical modifications in the RNA ribonucleoside backbone to make the chimeric RNA molecules more stable are well known in the art.

[0060] Different embodiments for chimeric genes or RNA molecules have been described above in relation to the provided methods for altering cellulose biosynthesis and can be applied *mutatis mutandis* to the embodiments relating to substances.

[0061] Chimeric genes or RNA may be provided to plant cells in a stable way, or transiently. Conveniently, stable provision of chimeric genes or RNA molecules may be achieved by integration of the chimeric genes into the genome of the cells of a plant. Methods for the introduction of chimeric genes into plants are well known in the art and include *Agrobacterium*-mediated transformation, particle gun delivery, microinjection, electroporation of intact cells, polyethylene glycol-mediated protoplast transformation, electroporation of protoplasts, liposome-mediated transformation, silicon-whiskers mediated transformation etc. The transformed cells obtained in this way may then be regenerated into mature fertile plants.

[0062] In another embodiment, the chimeric genes or chimeric RNA molecules of the invention may be provided on a DNA or RNA molecule capable of autonomously replicating in the cells of the plant, such as e.g. viral vectors. The chimeric gene or the RNA molecules of the invention may also be provided transiently to the cells of the plant.

[0063] It is also an object of the invention to provide plant cells and plants containing the chimeric genes or the RNA molecules according to the invention. Gametes, seeds, embryos, either zygotic or somatic, progeny or hybrids of plants comprising the chimeric genes of the present invention, which are produced by traditional breeding methods, are also included within the scope of the present invention.

[0064] The methods and means of the invention are suited for use in cotton plants, (both *Gossypium hirsutum* and *Gossypium barbadense*) including, but not limited to, plants such as

Coker 312, Coker310, Coker 5Acala SJ-5, GSC25110, FiberMax®819, FiberMax®832, FiberMax® 966, FiberMax® 958, FiberMax® 989, FiberMax® 5024 (and transgenic FiberMax® varieties exhibiting herbicide or insect-resistant traits) Siokra 1-3, T25, GSA75, Acala SJ2, Acala SJ4, Acala SJ5, Acala SJ-C1, Acala B1644, Acala B1654-26, Acala B1654-43, Acala B3991, Acala GC356, Acala GC510, Acala GAM1, Acala C1,Acala Royale, Acala Maxxa, Acala Prema, Acala B638, Acala B1810, Acala B2724, Acala B4894, Acala B5002, non Acala "picker" Siokra, "stripper" variety FC2017, Coker 315, STONEVILLE 506, STONEVILLE 825, DP50, DP61, DP90, DP77, DES119, McN235, HBX87, HBX191, HBX107, FC 3027, CHEMBRED A1, CHEMBRED A2, CHEMBRED A3, CHEMBRED A4, CHEMBRED B1, CHEMBRED B2, CHEMBRED B3, CHEMBRED C1, CHEMBRED C2, CHEMBRED C3, CHEMBRED C4, PAYMASTER 145, HS26, HS46, SICALA, PIMA S6 and ORO BLANCO PIMA.

[0065] The methods and means described herein may also be employed for other plant species such as hemp, jute, flax and woody plants, including but not limited to *Pinus spp., Populus spp., Picea spp., Eucalyptus spp.*, etc.

[0066] In another embodiment, a method for identifying allelic variations of the genes encoding proteins involved in cellulose biosynthesis in a population of different genotypes or varieties of a particular plant species, for example a fiber-producing plant species, which are correlated either alone or in combination with the quantity and/or quality of cellulose production, and fiber production is provided. This method comprises the following steps:

a) providing a population of different varieties or genotypes of a particular plant species or interbreeding plant species comprising different allelic forms of the nucleotide sequences encoding proteins comprising the amino acid sequences of SEQ ID No 5, 6, 7 or 8. The different allelic forms may be identified using the methods described elsewhere in this application. For example, a segregating population may be provided, wherein different combinations of the allelic variations of the genes encoding proteins involved in cellulose biosynthesis are present. Methods to produce segregating populations are well known in the art of plant breeding.

b) Determining parameters related to fiber production and/or cellulose biosynthesis for each individual of the population;

c) determining the presence of a particular allelic form of the nucleotide sequences encoding proteins comprising the amino acid sequences of SEQ ID No 5, 6, 7 or 8 for each individual of the population; and

d) correlating the occurrence of particular fiber or cellulose parameters with the presence of a particular allelic form of the mentioned nucleotide sequence or a particular combination of such allelic forms.

[0067] The resulting information will allow selecting those alleles which have the desired effect on cellulose biosynthesis or fiber production. The resulting information may be used to accelerate breeding programs, to isolate or create varieties with particular fiber or cellulose characteristics, or to accelerate backcross programs, by determining the presence or absence of allelic forms, using conventional molecular biology techniques. Methods for determining allelic forms in polyploid plants are known in the art and include e.g. Denaturing High-Performance Liquid Chromatography (DHPLC; Underhill et al. (1997) *Genome Research* 7:996-1005). It will be clear that not only the sequences of the alleles themselves can be used to determine their presence or absence during breeding or backcross programs, but also of the nucleotide sequences adjacent (e.g., immediately adjacent) and contiguous with the desired alleles, and which can only be separated from the allele by recombination during meiosis at low frequencies during meiosis.

[0068] As used herein "an interbreeding plant species" is a species which can be crossed with the fiber producing plant such as cotton (including using techniques such as hybridization etc.) and can produce progeny plants. Interbreeding plant species may include wild relatives of the fiber producing plants. Conventionally, for cotton plants reference is made to interbreeding for crosses between *G. barbadense* and *G. hirsutum* and to intrabreeding for crosses between two *G. barbadense* or two *G. hirsutum* parents.

[0069] The following non-limiting Examples describe method and means for modulating cellulose biosynthesis in fiber-producing plants. Unless stated otherwise in the Examples, all recombinant DNA techniques are carried out according to standard protocols as described in Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual,* Second Edition, Cold Spring Harbor Laboratory Press, NY and in Volumes 1 and 2 of Ausubel et al. (1994) *Current Protocols in Molecular Biology,* Current Protocols, USA. Standard materials and methods for plant molecular work are described in Plant Molecular Biology Labfax (1993) by R.D.D. Croy, jointly published by BIOS Scientific Publications Ltd (UK) and Blackwell Scientific Publications, UK.

Other references for standard molecular biology techniques include Sambrook and Russell (2001) *Molecular Cloning: A Laboratory Manual,* Third Edition, Cold Spring Harbor Laboratory Press, NY, Volumes I and II of Brown (1998) Molecular Biology LabFax, Second Edition, Academic Press (UK). Standard materials and methods for polymerase chain reactions can be found in Dieffenbach and Dveksler (1995) *PCR Primer: A Laboratory Manual,* Cold Spring Harbor Laboratory Press, and in McPherson at al. (2000) *PCR - Basics: From Background to Bench,* First Edition, Springer Verlag, Germany.

[0070] Throughout the description and Examples, reference is made to the following sequences:

SEQ ID No.1: *Arabidopsis* nucleotide sequence rsw2 (genomic; Accession number At5g4970).

SEQ ID No. 2: cotton nucleotide sequence rsw2 (cDNA)

SEQ ID No. 3: *Arabidopsis* nucleotide sequence rsw3 (genomic)

SEQ ID No. 4: cotton nucleotide sequence rsw3 (corresponding to the 3' end; cDNA)

SEQ ID No. 5: *Arabidopsis* amino acid sequence rsw2

SEQ ID No. 6: cotton amino acid sequence rsw2

SEQ ID No. 7: *Arabidopsis* amino acid sequence rsw3

SEQ ID No. 8: cotton amino acid sequence rsw3 (partial)

SEQ ID No. 9: *Arabidopsis* nucleotide sequence rsw2 (cDNA)

SEQ ID No. 10: oligonucleotide PCR primer (forward rsw2 cotton)

SEQ ID No. 11: oligonucleotide PCR primer (reverse rsw2 cotton)

SEQ ID No. 12: oligonucleotide PCR primer (forward LFY3)

SEQ ID No. 13: oligonucleotide PCR primer (reverse LFY3)

SEQ ID No. 14: oligonucleotide PCR primer (forward MBK5/α)

SEQ ID No. 15: oligonucleotide PCR primer (reverse MBK5/α)

SEQ ID No. 16: oligonucleotide PCR primer (At glucosidase II α forward)

SEQ ID No. 17: oligonucleotide PCR primer (At glucosidase II α reverse) 10

SEQ ID No. 18: oligonucleotide PCR primer (At glucosidase II β forward)

SEQ ID No. 19: oligonucleotide PCR primer (At glucosidase II β reverse)

SEQ ID No. 20: oligonucleotide PCR primer (forward primer to isolate genomic copy RSW3)

SEQ ID No. 21: oligonucleotide PCR primer (reverse primer to isolate genomic copy RSW3)

SEQ ID No. 22: oligonucleotide PCR primer (forward RWS3 homologue cotton)

SEQ ID No. 23: oligonucleotide PCR primer (reverse RSW3 homologue cotton).

[0071] Example 1. Isolation of a full length cDNA of the GhKOR gene (cotton gene corresponding to the rsw2 mutation in *Arabidopsis*).

[0072] The NCBI EST database has 7 ESTs from a *Gossypium arboreum* 7-10 dpa (days post anthesis) fiber library which show similarities to the sequence of AtKOR1. The sequences of five of the seven ESTs were identical. Alignment of the three different cotton ESTs against the AtKOR1 cDNA showed that cotton clone AW726657 contained the ATG start codon, and 47 bp of 5' untranslated region. Clone BE052640 spanned the middle region of the KOR gene and overlapped clone AW668085 which contained a TGA stop codon in the same position as that in AtKOR1 and 126 bp of 3' untranslated sequence. Translation of the ORF showed >80% amino acid sequence identity to regions of AtKOR1 protein. Primers designed to the 5' and 3' untranslated regions of the *G. arboreum* ESTs were used to amplify a 1.9 kb PCR product from an 18 dpa fiber cDNA library from the *G. hirsutum* cultivar Siokra 1-4. The forward primer was 5'- CCGCTCGAGCGGGCATTTTCCGCCCACTA-3' (SEQ ID No. 10) and the reverse primer 5'-CGGGATCCCGTCACACATGGACAGAAGAA-3' (SEQ ID No 11). A full length cDNA of the cotton KOR gene was generated by the PCR of a cotton cDNA library from 18 dpa fibers of *Gossypium hirsutum* and the products of several amplifications sequenced (SEQ ID No. 2). The cDNA encoded a protein (GhKOR) of 619 amino acids (SEQ ID No. 6) that was highly similar to LeCel3 (86% amino acid identity), AtKOR1 (82% amino acid identity) and BnCel16 (82% identity) (Fig. 1). All proteins shared: polarized targeting motifs involved in targeting AtKOR1 to the cell plate (Zuo et al., 2000); a putative transmembrane region near the N-terminus; four of the conserved residues potentially involved in catalysis (Asp-198, Asp-201, His-516 and E-555; Nicol et al., 1998) as part of the strong similarity to family 9 glycoside hydrolases; a C-terminal region rich in Pro and characteristic of membrane-bound members of the endo-1,4-β-D-glucanase family; 8 putative N-glycosylation sites (Asn-X-Ser/Thr) in the N-terminal domain predicted to be in the ER lumen during glycosylation. (An additional site present only in GhKOR (residues 14-16) would face the cytosol).

[0073] Example 2. Complementation of the Arabidopsis rsw2-1 mutant with GhKOR

[0074] The cotton PCR product encoding GhKOR was cloned behind the CaMV 35S promoter in the following way: the forward primer incorporated a XhoI site (underlined), and the reverse primer a BamHI site (underlined) which allowed the amplified 1.9 kb fragment to be ligated into the appropriate sites in vector pART7 (Gleave, 1992). This placed the cDNA in the sense orientation behind the cauliflower mosaic virus 35S promoter. The complete expression cassette was removed by digestion with NotI and cloned into the corresponding site in the binary vector pART27. The amplified product was sequenced to confirm its identity. This construct was introduced into *Agrobacterium tumefaciens* strain AGL1 and used to transform the rsw2-1 mutant and wild-type Columbia by floral dipping (Clough and Bent, 1998).

[0075] Kanamycin resistant transformants were selected on Hoagland's plates containing kanamycin (50 μg/ml) and timentin (100 μg/ml), transferred to vertical Hoagland's plates without selection agents and screened for root swelling after 2 days at 29°C. T2 seed was collected from ten individual T1 plants showing a wild-type phenotype and checked for inheritance of the complemented phenotype in the T2 generation. Photographs were taken of roots of T3 seedlings that were homozygous for kanamycin resistance and had been exposed to 29°C for 2 d. Other plants grown in pots at 21°C until the bolt was initiated had the bolt cut off before transfer to 29°C and the regenerated secondary bolts were photographed when mature. rsw2-1 has a single nucleotide change from Columbia in At5g49720 that replaces Gly-429 with Arg in AtKOR1 and provides a temperature-sensitive phenotype (Baskin et al.,1992; Lane et al., 2001). Plants were grown either in pots (1:1:1 mix of peat:compost:sand), or aseptically in Petri dishes (MS or Hoagland's medium with agar) (Burn et al., 2002a). Growth cabinets provided 100 μmol m^{-2} s^{-1} of continuous light at 21°C unless otherwise stated. Roots of the rsw2 mutant show temperature-sensitive radial swelling (Baskin et al., 1992) and stems show temperature-sensitive inhibition of elongation (Lane et al., 2001).

[0076] The roots of 63 out of 75 of the kanamycin-resistant T1 seedlings did not swell after 2 d at 29°C. The wild type phenotype was stably inherited into the T3 generation and roots (Fig 2A) and stems (Fig 2B) elongated normally at the restrictive temperature. Stem growth in T3 plants homozygous for kanamycin resistance was quantitatively indistinguishable from wild type. A gene was thus identified encoding a cotton homologue of AtKOR1 and it was shown that it can functionally replace the Arabidopsis gene in the rsw2-1 cellulose synthesis mutant.

[0077] This will involve GhKOR correcting defects in cytokinesis and cell elongation in *Arabidopsis* (Nicol et al., 1998; Zuo et al., 2000; Lane et al., 2001; Sato et al., 2001) as well as proper interaction with other elements of the cellulose synthesis machinery and/or products. Previous studies identified a cotton fiber protein immunologically related to LeCel3 (Peng et al., 2001) and indirect evidence implicated it in cellulose synthesis in vitro by cotton fiber membranes (Peng et al., 2002). The similarities to LeCel3, BnCel16 and AtKOR1 includes all major features of known functional significance and those, such as the Pro-rich C-terminus, which have no currently known function. The role of an endo-1,4-β-D-glucanase in cellulose synthesis is not clearly established but could involve severing a yet-to-crystallize glucan from a lipid-linked primer or donor (Williamson et al., 2001; Peng et al., 2002).

[0078] Example 3: Identification and isolation of the gene that has been mutated in rsw3 mutant of *Arabidopsis thaliana*.

[0079] The *rsw3* allele behaves as a single Mendelian recessive locus (Baskin et a!, 1992) and was identified by a map based strategy. The F2 progeny from crossing *rsw*3 with the visual marker line W9 linked RSW3 with *yi* on the lower arm of chromosome 5. An F2 population from crossing *rsw3* (Columbia background) with the Landsberg erecta ecotype was screened to give plants showing a root swelling phenotype. DNA was prepared from 2-3 rosette leaves per plant using the FastDNA kit (BIO 101, Carlsbad, CA) and mapping carried out using *LFY3* (forward primer 5'-GACGGCGTCTAGAAGATTC-3' (SEQ ID No. 12) , reverse 5'-TAACTTATCGGGCTTCTGC-3'; SEQ ID No. 13; cleavage with RsaI) and *MBK5/α* (forward 5'-CCCTCGCTTGGTACAAGGTAT-3' (SEQ ID No. 14) and reverse 5'-TCCTGATCCTCTCACCACGTA-3'(SEQ ID No. 15). Using the F2 from a cross to the Landsberg erecta ecotype, RSW3 was mapped at 6 cM from the *LFY3* locus (4 out of 70 chromosomes showing a cross over event) so positioning RSW3 between *yi* and *LFY3*. Analysis of a further 372 chromosomes identified one recombination event between *MBK5/α* and *rsw*3, a notional map distance of 0.27 cM. Several candidate genes in this region were sequenced in *rsw3*. One (At5g63840) on the P1 clone mgil9 (AB007646) encoded a putative catalytic subunit of glucosidase II and the *rsw*3 allele showed a T to C substitution predicted to replace Ser599 with Phe in the protein (nucleotide sequence of the wild type RSW3 gene is represented in SEQ ID No. 3, amino acid sequence of the encoded protein is represented in SEQ ID No. 7).

[0080] The RSW3 sequence is highly similar from about residue 150 onwards to sequences in the glucoside hydrolase family 31 (Henrissat, 1991; Henrissat and Bairoch, 1993). Monroe *et al* identified the *RSW3* glucosidase II gene through a search of *Arabidopsis* ESTs with homology to α-glucosidases and named it *Aglu-3* (Monroe *et al.,* 1999). Its protein product formed a clade with several glucosidase II enzymes whose catalytic activities were independently known. They all separated from apoplastic α-glycosidases of *Arabidopsis* with which Aglu-3/RSW3 shares only 8% sequence identity. Figure 4 shows the two signature motifs for the clade containing Aglu3/RSW3, which are believed to include catalytic and substrate binding residues. Aglu3/RSW3 contains all of the conserved residues within these motifs, as well as the proposed catalytic residues Asp512 and Asp617 (Frandsen and Svensson, 1998). Ser599, which is mutated in *rsw3,* is likely to be functionally significant since it is conserved in the homologous gene product from mouse (NP 032086), human (NP 055425), pig (AAB49757), slime mold (AAB18921), potato (P07391) and cotton (see below), and in the more distantly related apoplastic α-glucosidases encoded by the *Arabidopsis* genes *Aglu-1* and *Aglu-2* (Monroe *et al.,* 1999). The *Arabidopsis Aglu-3/RSW3* gene appears to be a single copy, spans 3.84 kb with 5 introns and encodes a predicted transcript of 2766 bp giving a predicted translation product of 104 kDa.

[0081] Recent biochemical (Trombetta *et al,* 1996) and genetic studies (D'Alessio *et al,* 1999; Pelletier *et al.*, 2000) suggest that native glucosidase II of mammals and yeast consists of a catalytic α-chain (to which Aglu-3/RSW3 is homologous) and a smaller non-catalytic ß-chain which retains the heterodimer in the ER. To determine if *Arabidopsis* contained an ortholog of the β-subunit, a BLAST search of the NCBI database was carried out with the mouse β-subunit. Unknown protein At5g56360 (protein MCD7.9 on the P1 clone MCD7 (AB009049) from chromosome 5) had 27% amino acid identity and 42% similarity to the mouse β-subunit. A closely related sequence (GenbankBAA88186) exists on chromosome 1 in rice but is annotated with a stop codon that truncates it after 496 residues. The conceptual translation of the adjacent 3' sequence on the PAC clone P0038F12 (AP000836) and reconsideration of proposed splice sites indicate the potential to encode a full length β-subunit that is very similar to the *Arabidopsis* gene product. The proposed sequence of the gene product is supported by an EST (AU030896) matching the proposed exons. Figure 5 therefore includes our suggestion for the full length rice protein. The *Arabidopsis,* rice, mouse and *Schizosaccharomyces pombe* sequences share: HDEL

ER-retention signals at the C-termini; predicted leader sequences at their N-termini; a cysteine-rich N-terminal region; a MHR (mannose-receptor homology region) (Munro, 2001) preceding the HDEL sequence at the C-terminus; a central region rich in acidic residues and flanked by regions giving high scores in programs ('Coils" and "Paircoil") predicting the likelihood of sequences forming coiled coils (Berger *et al.,* 1995; Lupas *et al,* 1991).

[0082] Munro (2001) links the MRH domain to carbohydrate recognition. It comprises a region of similarity to the cation-dependent mannose 6-phosphate receptor whose crystal structure is known. Critical conserved features (Figure 5) include the 6 Cys residues forming 3 disulphide bonds (although the *S. pombe* protein lacks cysteines 1 and 2), the substrate recognition loop between the cysteines 5 and 6 and the Y and R residues implicated in ligand binding (Roberts *et al.,* 1998). Interaction between mouse α and β subunits was mapped to the N-terminal 118 residues of the β-subunit, which are reasonably well conserved in all sequences, and to residues 273-400 (Arendt and Ostergaard, 2000) which are not. Figure 5 shows, however, that all sequences show a high percentage of acidic residues.

[0083] Expression of the genes encoding the α and β -subunits was analyzed using RT-PCR in the following way. RNA (Parcy *et al.,* 1994) was treated with RQl RNase-free DNase (Promega, Madison, WI) following the manufacturer's instructions. PCR primers were designed to the 3' end of the coding region of the α and ß-subunits of *Arabidopsis* glucosidase II :

α-forward 5'-CGTAGTGGTCTACTGGTTCAA-3' (SEQ ID No 16),

α-reverse 5'-TGAGCTGTGTCCCAAGAGGAT-3' (SEQ ID No. 17),

β-forward 5'-GGTGATGAGGATACCAGCGAT-3' (SEQ ID No. 18),

β-reverse 5'-CCCACTCCCTAACCGGAGTTT-3' (SEQ ID No. 19).

Each primer spanned an intron so differentiating RT-PCR products from genomic DNA and mRNA (724 bp versus 452 bp for the α-subunit, 996 versus 474 for the β-subunit). RT-PCR was carried out using the Gibco BRL Superscript one step RT-PCR kit, following the manufacturer's instructions and an RT-PCR cycle of 48°C 45min, 94°C 2min, (94°C/30sec, 54°C/1min, 68°C/2min)×45, 72°C -7 min. RT-PCR detected expression of the genes encoding the α and β-subunits in all tested tissues of *Arabidopsis* (Figure 6) but, under the conditions used, will not clearly indicate relative expression levels. The low numbers of ESTs in *Arabidopsis* (13 for the α- subunit, 4 for the β -subunit), suggest neither gene is highly expressed. (For comparison,

AtCesA1/RSW1, a glycosyltransferase implicated in cellulose synthesis, detects 40 ESTs in a similar search.)

[0084] Example 4: Complementation of the rsw3 mutation by a genomic copy of the *Arabidopsis* gene.

[0085] A genomic copy of the glucosidase II α-subunit including 830 bp of the promoter region was generated by PCR amplification of BAC F20A11 using the forward primer 5'-CCGCTCGAGCGGTTTCACTCACAACTGTGGTCTCT-3' (SEQ ID No. 20) and the reverse primer 5'-CCGCTCGAGCGGTCTCCTAAGTCCTAACCCCATA-3'(SEQ ID No. 21). Both primers included a *Xho*I site (underlined) which allowed the amplified 5.8 kb fragment to be ligated into the *Sal*I site in the binary vector pBin19. The amplified product showed a single base pair change (C to T) from the genomic sequence. This substituted Leu for Ser 142, a residue that is conserved in potato but not in other species (Figure 4) and did not impair the ability of the fragment to complement *rsw3.* The construct was introduced into *Agrobacterium tumefaciens* strain AGL1 and used to transform the *rsw3* mutant by floral dipping (Clough and Bent, 1998). Kanamycin- resistant transformants were selected at 21°C on Hoaglands's plates containing kanamycin (50 μg ml^{-1}) and timentin (100 μg ml^{-1}). Healthy seedlings were transferred to vertical Hoagland's plates and placed at 30°C for 2 days to screen for root swelling. Kanamycin resistant T1 progeny had wild-type roots when grown for 5 days at 21°C followed by 2 days at 30°C (Figure 3a). The inflorescence phenotype (see later) was also complemented.

[0086] A second line of evidence was provided by crosses between *rsw3* and the tagged mutant SGT5691 (Parinov *et al.,* 1999), which contains a Ds element in the first exon of the gene encoding the putative glycosidase II enzyme. It presumably represents a null allele and the mutation is homozygous lethal so hemizygous plants, which appear wild type, were used for crossing. The NPTII gene present on the Ds element confers kanamycin resistance to F1 plants receiving the tagged allele from SGT5691. Roots of all kanamycin-resistant F1 seedlings (containing a null allele and a temperature-sensitive allele) appeared wild-type at 21°C but swelled at 30°C (Figure 3b). This confirms that the Ds insertion mutant and the EMS generated mutant *rsw3* are allelic and that glucosidase II defects cause radial swelling.

[0087] Example 5: Observations on other phenotypes associated with the *rsw3* mutation in *Arabidopsis*.

[0088] *rsw3* grows like wild type at its permissive temperature of 21° C and the seedling root swells when transferred to 30°C. The bulging cells on the root (Baskin *et al.*, 1992) are often at the base of root hairs suggesting a role for RSW3 in the early stages of root- hair development. The swollen primary root only resumes elongation if returned to the permissive temperature within 48 h but the root continues to generate laterals (Figure 7a). The laterals - whose primordia were not visible when the transfer to 31°C was made- elongate for several mm before they in turn swell and stop growing. The root system of mature vegetative plants is consequently short and very highly branched (Figure 7b). The double cellulose-defective mutant *rsw1-rsw3* showed only a slightly swollen root tip after 24 h at the restrictive temperature but since any longer period at the high temperature led to death, swelling was probably already curtailed after 24 hours at the restrictive temperature.

[0089] The phenotype in dark-grown hypocotyls is much weaker in *rsw3* than in *rsw1-1* and *rsw2-1* and the phenotype in *rsw1-1rsw3* is weaker than *rsw1-1rsw2-1* (Figure 7c). Rosette growth of *rsw3* in the light is strongly suppressed and many minute leaves are packed in a dense mat in which regular phyllotaxis cannot be recognized (Figure 7d-f). The complex pavement cell shape in wild-type leaves (Figure 7g) is simplified in *rsw3,* stomata protrude from the leaf surface and some trichomes appear to burst (Figure 5h). Several of the crowded rosettes initiated minute inflorescences (Figure 7d) although these appear much later than wild-type inflorescences (28.6 ± 0.5 days versus 15.5 ± 0.17 days for agar grown plants; mean ± SE, n = 98 for *rsw3*, n = 45 for wild type). The few flowers on the minute *rsw3* inflorescences were essentially full-sized although anther filaments, gynoecium and sepals were slightly shortened and buds opened prematurely before the stigma was receptive (similar to the buds from soil grown *rsw3* plants shown in Figure 8e, f which are discussed below).

[0090] To investigate the direct effects of the mutation on stem growth, wild-type and *rsw3* were grown at 21°C on soil so that subsequent inflorescence development would not be limited by a small rosette supplying little photosynthate. Rosettes of *rsw3* were very similar to wild type under these conditions and reproductive growth began at the normal time.

[0091] Primary bolts were cut off and regrowth of secondary bolts followed at either 21°C or 30°C (Figure 6a, b). Regrowth followed a slightly S-shaped curve with *rsw3* and rsw*1-1* at 21°C

showing statistically insignificant reductions in growth rate and final height relative to wild type. *Rsw1-1rsw3* showed a clear reduction in rate and final height. At 30°C, however, the *rsw3* growth rate was similar to wild type for a few days but elongation stopped by about day 5 whereas it continued in wild type until day 16 and even longer in *rsw1-1* (Figure 8b). *rsw1-1rsw2* (Lane *et al.*, 2001) failed to regenerate secondary bolts at 30°C and *rsw1-1rsw3* only grew to about 35 mm (Figure 8b) and produced few flowers and no seed.

[0092] Measurements of daily stem growth increments and the lengths of epidermal cells, which had left the elongation zone when the bolts were about half grown (Table 1), were made. This allowed estimation of cell flux (the number of cells leaving the elongation zone day^{-1}) at that time since daily growth increment = cell length × cell flux. There was no significant reduction in either cell flux or cell length of *rsw3* growing at 21°C. The *rsw1-1rsw3* constitutive phenotype at 21°C was entirely due to a reduction in cell length. At 30°C, *rsw1-1* showed a 57% reduction in cell length and a 35% reduction in cell flux relative to wild type.

[0093] Analyses of this type require that the plant is in a near steady state with respect to growth rate, length of the elongation zone etc. Conditions, however, are far from steady state when elongation is rapidly slowing in *rsw3* and *rsw1-1rsw3* so that accurate deductions of cell flux for those genotypes are precluded. To get at least an idea of how cell length was behaving when growth was slowing, we measured cell lengths at a height of about 80 mm on the *rsw3* stem. (Figure 8b shows that when these cells left the elongation zone, the stem would have been near the end of its growth phase since total plant height at that time would have exceeded 80 mm by the length of the growth zone at that time; 40 mm in wild type according to Fukaki *et al.*, 1996). The cells in *rsw3* were, even then, only slightly shorter than wild type (Table 1) suggesting that falling cell production rates were probably more important than reduced cell expansion in slowing stem elongation. In contrast, when we sampled the *rsw1-1rsw3* stem at 30 mm for cells maturing when its elongation was slowing (Figure 8b), cell length was reduced by 57% (Table 1). This is consistent with the presence of *rsw1-1* in the double mutant tilting the balance strongly towards reduced cell length.

[0094] These conclusions regarding cell division and cell expansion were checked in a simpler system by using cryo-scanning electron microscopy to examine stamen filaments in flowers showing receptive stigmas (Table 2). The results were similar: *rsw3* plants again showed a greater percentage reduction in cell number than in cell length and the double mutant *rsw1-1rsw3*

showed a further reduction in cell length without an additional reduction in cell number. *Rsw1-1* showed a much greater reduction in cell length than in cell number (Table 2). Stems of both wild type and *rsw3* regenerating at 30°C reached approximately the same height before initiating their first flower even though their final heights would be very different (Figure 8b). Wild-type stems generated about 27 well spaced flowers before elongation ceased but *rsw3* produced only about 6 closely spaced flowers before elongation ended leaving a cluster of flowers (Figure 8c, d). *rsw3* flower buds opened precociously before the stigma was receptive (Figure 8e, f).

[0095] Few flowers and no seed formed on the minute bolts of *rsw3* plants grown continuously at their restrictive temperature (Figure 7d). Even flowers on the much larger bolts formed at 31°C on plants which had completed vegetative growth at 21°C (Figure 8d, f) also set very little seed. That seed (Figure 8g, h) was shrunken (probably because of reduced accumulation of seed storage proteins; Boisson *et al.,* 2001), its surface lacked the regular cellular structure of wild type grown at 30°C or of *rsw3* grown at 21°C and it showed very little secreted mucilage after imbibition (Figure 8i-n). Reduced mucilage secretion was not typical of cellulose-deficient mutants: *rsw1-1* (defective in the CesA1glycosyltransferase; Figure 8 k, l), and *rsw2-1* (defective in the KOR endo-l,4 β glucanase) had normal mucilage coats.

[0096] To isolate effects on the haploid stages of pollen and ovule development from effects on the diploid stages, we examined seed set in the hemizygous Ds-mutant SGT5691 (a presumed null allele in the glucosidase II catalytic subunit). Seed set by self-fertilization segregates 147 kanamycin-resistant individuals to 153 sensitive individuals. A ratio less than the 2:1 expected for a dominant, homozygous lethal allele shows that the null allele affects post-meiotic development of pollen and/or ovules. We separated the effects on the male and female pathways by reciprocal crosses between the hemizygous tagged mutant and Landsberg erecta (the appropriate wild type for this mutant). Kanamycin-resistant and sensitive plants will segregate 1:1 if pollen or ovule development is unaffected with lower ratios if the null allele reduces pollen or ovule fertility. Pollen from the Ds-tagged mutant gave a segregation ratio of 1:16 (6 resistant:94 sensitive individuals) indicating a 94% reduction (relative to wild type) in the ability of Ds-tagged pollen to set viable seed. This compared with a 41% reduction when Ds-tagged ovules were crossed to wild type pollen (ratio of 1:1.7, 37:63 individuals). The null allele of glucosidase II therefore affects the haploid stages of pollen development much more severely than it affects post-meiotic ovules.

[0097] Roots of 7 day old seedlings of *rsw3* grown at 31° C contain only 51 % of the wild-type cellulose (expressed mg-1 tissue dry weight), a comparable figure to that resulting from single amino acid substitutions in the CesA1 glycosyltransferase *(rsw1-1)* and the KOR endo-1,4-β-glucanase *(rsw2-1)* (Peng *et al.,* 2000). The morphological changes indicate that all three genes are needed to make cellulose in primary cell walls.

[0098] Production of Golgi-derived non-cellulosic polysaccharides changes little in *rsw3* seedlings (Peng *et al.,* 2000). The selectivity for cellulose production is comparable to that seen with a defect in glucosidase I (Gillmor *et al.,* 2002), the enzyme generating the initial substrate for glucosidase II processing. It exceeds the selectivity seen in the embryo-lethal *cyt1* mutants of *Arabidopsis* (defective in mannose-1-phosphate guanylyltransferase) (Lukowitz *et al.,* 2001) and in potatoes with MALl (encoding a glucosidase II α-subunit) down-regulated by antisense (Taylor *et al.,* 2000a) where complex changes occur in non-cellulosic polysaccharides and lignin. We therefore conclude that cellulose synthesis is often much more sensitive to N-glycan processing defects than is the synthesis of non-cellulosic polysaccharides in the Golgi.

[0099] Secretion of Golgi-derived seed mucilage is strongly reduced in *rsw3* but not in rsw*1-1* or *rsw2-1.* Mucilage could be produced but retained intracellularly (perhaps because of structural changes resulting from cellulose deficiency), or mucilage production itself could be reduced. Many developmental blocks reduce mucilage production (Western *et al.,* 2001; Western *et al.,* 2000) but we cannot yet exclude the possibility that *rsw3* has defective processing of Golgi enzymes required to make the particular non-cellulosic polysaccharides making up the mucilage.

[0100] Cell numbers and sizes in stamen filaments indicate that *rsw3* affects cell division more strongly than cell expansion. The cell length data for the stem are consistent with this finding. A strong effect of *rsw3* on cell division may explain why its phenotype is rather weak in dark grown hypocotyls which lack cell division (Gendreau *et al.,* 1997). In more strongly affecting cell division than cell expansion, *rsw3* resembles *rsw2-1* (Burn *et al,* 2002) rather than *rsw1-1* (Burn *et al.,* 2002) or plants carrying antisense constructs to *RSW1/CesA1* or *CesA3* (Burn *et al.,* 2002) which are more severely affected in cell length. (Although *CesA1* changes have little impact on division rates, *CesA1* is probably expressed in dividing root cells since they show changes in wall ultrastructure (Sugimoto *et al.,* 2001) and swell (Baskin et *al.,* 1992; Beemster and Baskin, 1998) when *rsw1-1* is at its restrictive temperature.)

[0101] Although it is clear that cellulose biosynthesis is impaired in the *rsw3*, the mechanism by which rsw3 affects cellulose synthesis is not yet clear. As noted in relation to a glucosidase I mutation (Boisson *et al.,* 2001), the minimal phenotype shown by a mutant which cannot assemble mature N-linked glycans in the Golgi (von Schaewen *etal,* 1993) indicates that a lack of mature N-linked glycans on critical proteins will not cause the strong phenotype seen with a glycosidase II defect. Reduced rates of production of $Glc_1Man_9GlcNAc_2$ and $Man_9GlcNAc_2$ would probably slow both the formation and dissociation of the glycoprotein/chaperone complex creating a bottleneck that may in time reduce the steady state levels of glycoproteins at sites further along the secretory pathway. Because glycoproteins participate in many plant processes, it is not obvious why cellulose synthesis should be much more sensitive to processing defects in the ER than, for example, synthesis of non-cellulosic polysaccharides.

[0102] Gillmor *et al.* (2002) argued that CesA proteins are not glycosylated when they did not detect a mobility shift on SDS -PAGE in *knopf* (deficient in glucosidase I) or alter N-glycosidase F treatment and when they did not see in *knopf* a change in CesA abundance that was visible by unquantified immunostaining. The KOR endo-1,4-β-glucanase is a better candidate. A soluble fragment of the *Brassica napus* ortholog of KOR is heavily N- glycosylated when expressed heterologously in *Pichia pastoris* and the N-glycan is required for *in vitro* activity (Molhoj *et al.,* 2001). Further evidence consistent with KOR being a target can be drawn from the *rsw3* and *rsw2-1* phenotypes affecting cell division more than cell expansion whereas the *rsw1-1* phenotype shows the reverse.

[0103] The *rsw1-1* and *rsw2-1* mutations affect genes encoding plasma membrane enzymes that are probably directly involved in cellulose synthesis so that changed enzyme performance at the restrictive temperature will rapidly impact on cellulose synthesis. *rsw3,* in contrast, encodes a processing enzyme in the ER whose changed performance will reduce cellulose synthesis only when it restricts the supply of properly folded glycoproteins to the site of cellulose synthesis. The different time courses for the onset of a visible phenotype when the three mutants are transferred to the higher temperature plausibly reflect these different modes of action. Radial swelling starts slowly in *rsw3* (latency>24 h compared to < 12 h in *rsw1-1* and *rsw2-1*) and the high temperature actually accelerates root elongation during the first 12 h, albeit by less than in wild type (Baskin *et al.,* 1992).

[0104] Elongation of *rsw1-1* or *rsw2 -1,* in contrast, falls during the first 12 h, roots swell strongly and *rsw1-1* shows changed wall ultrastructure within 4 h (Sugimoto *et al.,* 2001).

[0105] It has been shown that *rsw3* is mutated in a gene encoding a putative glycosidase II α-subunit, identified a putative β-subunit encoded by two plant genomes and shown that many aspects of the *rsw3* phenotype flow from reduced cellulose synthesis in primary walls. Cell division seems more strongly affected than cell expansion indicating that the KOR endo-l,4-β-glucanase, where mutations also strongly affect cell division, may be the glycoprotein affected by the processing defect. In addition to its role in cellulose synthesis, a temperature-sensitive allele of glucosidase II will contribute to studies of N-glycosylation and quality control in the ER and in establishing its links to other developmental and physiological processes.

[0106] Example 6: Isolation of a (partial) cDNA corresponding to *RSW3* from cotton

[0107] A dbEST search using the sequence of *RSW3* as query, identified a *Gossypium arboreum* cDNA with 833 bp of high quality sequence. Primers designed from the EST were used to amplify a 700 bp product form a library of 18 dpa fibers of G. hirsutum cDNA using the following primers:

Cot-rsw3f 5'-CGGGATGAAGAGGATGTAGAG 3' (SEQ ID No. 22)

Cot-rsq3r 5'-GAACCCCTGAGATGATCCCAA 3' (SEQ ID No. 23)

[0108] The PCR product was used as a probe to identify longer cDNAs. 5 putative clones were identified and 2 were sequenced. The three clones overlapped and the sequence of cDNA of the cotton RSW3 homolog was assembled (SEQ ID No. 4). The region encoding the N-terminus is missing.

[0109] Example 7: Expression of *RSW2/RSW3* chimeric genes in cotton

[0110] cDNAs corresponding to *RSW2 or RSW3*, isolated from *Arabidopsis* or cotton are operably linked to a promoter such as the expansin promoter and a 3' end region involved in transcription termination and polyadenylation.

[0111] Further, about 100 bp long fragments selected from the *RSW2 or RSW3* genes isolated from *Arabidopsis* or cotton are cloned in inverted repeat under the control of a promoter such as the CaMV35S promoter.

[0112] The chimeric genes are introduced into a T-DNA vector comprising further a selectable marker gene, and the resulting T-DNA vectors are introduced into *Agrobacterium tumefaciens* strains containing a helper Ti-plasmid. Transgenic cotton plants are obtained using these Agrobacterium strains.

[0113] Plants expressing copies of the different transgenes are analyzed further for cell wall components, including cellulose, non-crystalline β-1,4 glucan polymer, starch and carbohydrate content as described in WO 98/00549.

[0114] Table 1. Analysis of the rate of stem elongation in terms of cell length and, where near steady growth rates occurred, cell flux (number of cells day^{-1} leaving the elongation zone). Results are given as mean + SE for n=5. Statistically significant differences from wild type using the Student's T-test are indicated (* = $p < 0.05$;** = $p < 0.01$; *** = $p < 0.001$).

		Growth rate (mm day^{-1})	Cell flux (day^{-1})	Cell length (μm)
21°C	Columbia	38.7 ± 1.0	101 ± 3.5	384 ±4.0
	rsw3	38.4 ±1.4	95.9 ± 4.6	402 ±7.0
	rsw1	38.9 ± 1.6	102 ± 6.9	382 ± 9.8
	rsw1rsw3	30.2 ± 1.9**	100 ±7.6	299 ± 8.4**
30°C	Columbia	53.8 ±1.2	133 ± 2.7	404 ± 3.2
	rsw3	41.8±3.1**		378±22
	rsw1	15.2± 1.4***	87.2 ± 7.0**	174 ± 5.8***
	rswlrsw3	13.6± 1.8***		173 ±15***

[0115] Table 2. Cell length and number in mature stamen filaments grown at 30°C. Results are given as mean + SE for n >7. Statistically significant differences from wild type using the Student's T-test are indicated (*= p <0.05; ** = p< 0.01; *** p = <0.001).

	Total length (μm)	Cell number	Cell length (μm)
Columbia	2407 _ 38	17.0 _ 1.0	152.7 _ 6.2
rsw3	1458 _ 52***	11.4 _ 0.3***	127.0 _ 0.1**
rsw1-1	1050 _ 57***	15.0 _ 0.4	72.7 _ 9.8***
rsw1-1rsw3	415 _ 41***	12.4 _ 0.5***	29.4 _ 2.1***

[0116] References

Arioli et al.. (1998) *Science* 279: 717-720.

Arendt et al. (2000) *Glycobiology* **10:** 487-492.

Baskin et al.. *Aust. J. Plant Phys.* 19: 427-437.

Beemster et al. (1998) *Plant Physiol*, 116: 1515-1526.

Berger et al. (1995) *Proc. Nat Acad. Sci. USA* **92:** 8259-8263.

Boisson et al. (2001) *EMBO J.*, 20: 1010-1019.

Brada et al. (1984) *Eur. J. Biochem.*, 141: 149-156.

Brummell et al.*Proc. Nat. Acad. Sci. USA* 94: 4794-4799.

Burn et al (2002a) *Plant Physiol.* 129: 797-807.

Clough et al. (1998) *Plant J.* 16: 735-743.

D'Alessio et al. (1999) *J. Biol. Chem.*, 274: 25899-25905.

Desprez, et al. (2002) *Plant Physiol.* 128: 482-490.

Fagard et al. (2000) *Plant Cell* 12: 2409-2424.

Frandsen et al. (1998) *Plant Mol. Biol.* 37: 1-13

Fukaki et al. (1996) *Plant Physiol.* 110, 933-943.

Gendreau et al. (1997) *Plant Physiol.*, 114: 295-305.

Gillmor et al. (2002) *J. Cell Biol.* 156: 1003-1013.

Gleave, A.P. (1992) *Plant Mol. Biol.* 20: 1203-1207.

Helenius et al. (2001) *Science* 291: 2364-2369.

Henrissat (1991) *Biochem. J.* 280: 309-316

Henrissat and Bairoch (1993) *Biochem. J.* 293: 781-788.

Hino and Rothman (1985) *Biochemistry* 24: 800-805.

Kimura et al.(1999) *Plant Cell* 11: 2075-2085.

Lane et al.(2001) *Plant Physiol.* 126: 278-288.

Lukowitz et al. (2001) *Proc. Nat. Acad. Sci. USA* 98: 2262-2267.

Lupas et al. (2001) *Plant Physiol.* 127: 674-684.

Mølhøj et al. (2001) *Plant Physiol.* 127: 674-684.

Monroe et al. (1999) *Plant Physiol.* 119: 385-397.

Munro (2001) Curr. Biol. 11: R499-501.

Murashige and Skoog (1962) *Phys. Plant* 15: 473-497.

Nicol et al. (1998) *EMBO J.* 17: 5563-5576.

Pagant et al. (2002) *Plant Cell* 14: 2001-2013.

Parcy et al. (1994) *Plant Cell* 6: 1567-1582

Parinov et al. (1999) *Plant Cell* 11: 2263-2270.

Pelletier et al. (2000) *Glycobiology* 10: 815-827.

Peng et al. (2000) *Planta* 211: 406-414.

Peng et al. (2002) *Science* 295: 147-150.

Peng et al. (2001) *Plant Physiol.* 126: 981-992.

Roberts et al. (1998) *Cell* 93: 639-648.

Sato et al. (2001) *Plant Cell Physiol* 42: 251-263.

Scheible et al. (2001) *Proc. Nat. Acad. Sci. USA* 98: 10079-10084.

Silk et al. (1989) *Plant Physiol.* 90: 708-713.

Sugimoto et al. (2001) *Protoplasma* 215: 172-183.

Taylor et al. (2000a) *Plant J.* 24: 305-316

Taylor et al. (2000) *Plant Cell* 12: 2529-2539.

Taylor et al. (1999) *Plant Cell* 11: 769-780.

Treml et al. (2000) *Glycobiology* 10: 493-502.

Trombetta et al. (2001) *Biochemistry* 40: 10717-10722.

Trombetta et al. (1996) *J. Biol. Chem.* 271: 27509-27516.

Vitale (2001) *Plant Cell* 13: 1260-1262

von Schaewen et al. (1993) *Plant Physiol* 102: 1109-1118.

Western et al. (2001) *Plant Physiol* 127: 998-1011.

Western et al. (2000) *Plant Physiol* 122: 345-355.

Williamson et al. (2001a) *Protoplasma* 215: 116-127.

Williamson et al. (2001) *Cell. Molec. Life Sci.* 58: 1475-1490.

Zuo et al. (2000) *Plant Cell* 12: 1137-1152.